INTRODUCTION
Sociality and the human mind

Julian Kiverstein

Introduction

The idea of humans as by nature social and political and animals can be traced back to Aristotle and was given a modern inflection by Hegel and Marx. These philosophers took sociality to be built into our very being. It is what defines us as human, distinguishing us from other animals. Yet how did we come to be this way? Chimpanzees live in groups, form alliances and they use simple tools such as sticks to fish for termites and stones to break open nuts. They have the rudiments of a social and political life. However, a gulf seems to separate their social life from ours. Humans created political and economic institutions. We developed elaborate systems of knowledge of our own history, and of the natural and physical world. We enrich our lives with artistic objects, each the product of a long prior history of making which these objects reflect back at us. We identify ourselves with groups and engage in the complex rituals and practices of those groups. We feel guilt, shame and pride when a member of our own group does something noteworthy. By contrast, we instinctively fear and often despise members of out-groups. Is there something in human psychology that might explain how people joined forces to become not just you and me but we?

This handbook comprises 32 original chapters and brings together perspectives from diverse disciplines from diverse areas of research in philosophy, comparative and developmental psychology, evolutionary anthropology, cognitive neuroscience and behavioural economics. Why should a handbook in philosophy draw upon the resources from these diverse disciplines?

The social life of humans seems on the face of it to make humans special. It seems to be among the sources of human uniqueness, but are humans really so special? Philosophers are not well-placed to answer this question on their own. It is a question that calls for careful comparison of the cognitive profile of humans with that of other species, in particular our closest evolutionary ancestors. While their social lives are in important respects different from our own, they are also in many ways very similar. These similarities might lead us to question the very idea of human uniqueness. Could the idea that we humans are somehow special be a product of anthropocentric philosophical preconceptions (see Barrett, ch.1)? At the very least, the uniqueness of humans isn’t something that should be taken for granted. It is a claim we ought to first carefully and critically scrutinise, bringing to bear all of the available resources from empirical and philosophical research.
The unrivalled complexity of human social life is a datum that stands in need of explanation. People invented science, art, religion and politics. No other animal has the sophisticated systems of symbolic communication and reasoning that humans have created. This presents us with something of a puzzle, explored by many of the authors in this volume in different ways. What is it about humans that enabled us to construct a social reality of unrivalled complexity? Is there something about the human mind that explains how we came to have social lives organised around social conventions, norms and institutions?

This volume is organised around the assumption that a broadly naturalistic answer to this question can be given. Humans are cultural beings and our lives undoubtedly have taken on very different forms within different cultural and historical contexts (Geertz 1977). Naturalists argue, however, that an explanation can be given of how history, culture and the social can also be conceived of as material processes. In my view there need be no competition between this type of naturalistic enquiry and theorising in the social sciences and the humanities. Any natural science of human behaviour must study human action within the contexts in which it occurs, taking into account the subjective perspective and points of view from which human agents act. Social scientists provide descriptions and interpretations of what it is like to be a social actor belonging to a particular society and culture. The explanations of the natural sciences fail to deliver the “thick” understanding of the particularities of human action we get from the internal standpoint of the social sciences (see Boutel & Lewens, ch.3). Naturalists argue, however, that cultural and social life stand in need of explanation if we are to avoid a split between bodies as products of biological evolution, and the human mind as a product of culture. It is such a naturalistic project that the contributors to this volume are engaged in.

Why speak of a social mind? I found three broad answers to this question in reading the chapters in this volume, and no doubt there are many more. The first takes its lead from evolutionary psychology. The human mind is a social mind because it is made up of specialised mechanisms for dealing with the problems of living in large social groups such as cheater detection, mate choice, disease avoidance, coalition formation and so on. Perhaps the most influential articulation of this theory can be found in Humphrey (1976). Humphrey sketched an evolutionary scenario in which the driving force behind the evolution of human intelligence was the elaborate social interactions among primates. Primates live in large groups consisting of many individuals who recognise each other, and are able to keep track of affiliations. They use this information to compete with each other for available resources, often entering into impressive acts of tactical deception (Byrne & Whiten 1989). Humphrey argued that in such an environment it will have been important to keep track of the mental states of one’s cohorts. The social environment of primates may thus have created selection pressures for capacities for meta-representation or thinking about thinking. The capacity for tracking the mental states of others so as to predict and manipulate their behaviour was an adaptation to a social environment containing agents that always have a stake in behaving deceptively. Sperber argues that this capacity for meta-representation makes human cognition especially powerful, enabling everything from cooperative communication to teaching and dialogical argument (Sperber 1996, 2000).

A second reason for thinking of the human mind as essentially social comes from the joint and collective forms of intentionality that distinguish the human mind from that of other animals. In recent years, a number of evolutionary hypotheses have been developed arguing that people are born default cooperators (Tomasello 2009, 2014; Sterelny 2003, 2012). Humans are distinguished from other social animals by the extent to which we enter into collaborative and cooperative social interactions. When chimpanzees work together in groups they tend to do so in what Raimo Tuomela (2007) has called the “I-mode” in pursuit of their own individual
goals. Our human foraging ancestors (hominins from around 400–200,000 years ago), by contrast, worked together collaboratively, and shared the fruits of their efforts with each other based on norms of fairness. This difference in social lifestyle may have created selection pressures for capacities for group-mindedness that didn’t arise for other primates. Youn infants enter the social world with capacities for shared intentionality and joint attentional engagement that are not present to the same degree in great apes. These capacities form the basis for people to coordinate with each other in ways that set the stage for agreeing upon conventions and norms.5

The third perspective takes the human mind to be the result of a co-evolutionary process in which biological evolution joins forces with cultural learning.6 Cultural learning allows a group to preserve technological innovations and skills whilst also building upon and extending those skills. People are able to transmit, preserve and elaborate on skills and knowledge across generations (Tomasello 1999). Children come to resemble their parents not only because of the genes they have inherited, but also because of the informational resources they have inherited from their parents (Sterelny 2003, 2012; Menary 2007). The minds of children grow in a learning environment structured by the skill base acquired from previous generations. This skill base includes facility with increasingly elaborate technical systems and public systems of representation. These tools and techniques have transformed the cognitive profile of humans, allowing us to develop and acquire types of thinking and problem solving that would be impossible had these systems not been invented.7

The three naturalistic approaches to understanding the sociality of the human mind are explored and developed in detail across the five parts of this handbook. Part 1 asks what might have happened over the course of human history that led to humans acquiring the abilities for the construction of a social world. Each of the three conceptions of the social mind I have just outlined make their first appearance in this section of the handbook. Part 2 zooms in on questions concerning the development of social cognition without which there would arguably be no social learning and no cooperative communication. The essays in part 3 delve into the nature of social and moral norms. The essays in this section draw their naturalistic inspiration from findings in cognitive neuroscience concerning social learning, and show how these findings can help us to understand human moral behaviour. Part 4 investigates questions raised by human capacities for joint and collective forms of thought and action, and the differences in thinking among a plurality of agents that share a perspective on the world. Part 5 brings the handbook to a close, and begins by looking at reasons for thinking that certain forms of selfhood and mindedness are constitutively social. The remainder of my introduction is organised around a brief synopsis of each chapter identifying and mapping some themes and topics that are central in the philosophy of the social mind. It should also help the reader to gain an overall view of the landscape covered in this volume.

1. The evolution of the social mind

The chapters in this first section explore the hypothesis that the human mind is social because human psychology has evolved to include cognitive and emotional capacities dedicated to social life. One hypothesis common to these chapters is that human cognition is distinguished from that of other primates by capacities for dealing with the social world. Some have argued these capacities took the form of sophisticated capacities for reasoning about the mental states of others, their beliefs, preferences and intentions. Others stress the importance of capacities for social learning which will have allowed for individuals to learn all kinds of new things from others in their culture including the use of artefacts and symbolic forms of communication. Evolutionary history may also have left its mark on human cognition in a number of social
domains from the way people think about social groups to the capacity for agent-neutral thinking that allowed for the emergence of conventions, norms and institutions.

In the opening chapter of this section, comparative psychologist Louise Barrett reviews the evidence for what has come to be known as the “social intelligence” or “social brain” hypothesis and finds the evidence wanting. She argues that the inference that is made from brain size to the cognitive capacities for dealing with the demands of social life is based on anthropocentric assumptions. The social intelligence hypothesis assumes that the social life of non-human primates (and many other species) calls for the same types of folk psychological abilities as are found in humans. This is anthropocentric insofar as it leads to the behavioural data being interpreted based on a projection of cognitive capacities found in humans onto other species.

The social intelligence hypothesis supports a theory of the social mind as a set of adaptations to social life characterised by competition for access to resources and deception. An alternative strand of thinking within evolutionary psychology views the social mind of humans as the result of a co-evolutionary process in which biological evolution joins forces with cultural learning. There are at least two forms that this co-evolution of the human mind and the social environment has taken. The first is discussed in chapter 2 by Richard Moore and concerns the co-evolution of social and technical skills. The second is discussed in chapter 3 by Adrian Boutel and Tim Lewens and concerns cultural evolution as a medium for the inheritance of survival enhancing behavioural traits.

Moore begins chapter 2 by explaining how the emergence of cumulative culture couldn’t be explained by individual learning alone. Moore argues that cumulative culture wouldn’t exist were it not for cultural forms of learning. He discusses two varieties of cultural learning: imitation and pedagogy. Both are high-fidelity modes of information transmission. In imitation an agent learns how to reproduce not just the outcome of an observed action but also the precise technique for bringing about this outcome. This type of learning guarantees that the precise details of a craft are preserved across generations. Moore defines “pedagogy” in terms of a teacher providing verbal instruction with the intention of helping a student to acquire knowledge or skills that are in some way important to a community. The student is thereby initiated into community and becomes responsible for maintaining and expanding the knowledge of this community.

In chapter 3 Adrian Boutel and Tim Lewens review theories of cultural evolution assessing to what extent these theories support a theory of the human mind as a social phenomenon. Models of cultural evolution seek to explain cultural and social change in Darwinian terms. Large-scale processes of cultural change are explained by aggregates of many smaller-scale events taking place in the lives of individuals. Cultural evolution is a theory of the social mind because it takes biases and dispositions that determine from whom we learn, and how we learn, to be built into the very fabric of the human mind. Capacities for social learning are genetically inherited adaptations that conferred reproductive advantages on individuals. These capacities tended to get us information that was selectively advantageous, and this in turn helped to spread the genes responsible for the capacities for social learning. At the same, Boutel and Lewens are keen to stress that cultural evolution offers a methodological individualist account of the social. The persistence and spread of cultural variants through populations is explained by the social acquisition of information by individuals.

In chapter 4 Richard Menary and Alexander James Gillett offer an alternative perspective on how cumulative culture may have contributed to the evolution of the human mind. Early humans developed skills that enabled them to adapt to a wide variety of different environments thereby creating a selection pressure for the extraordinary learning-dependent plasticity found in the human brain. Furthermore, humans go through an extraordinarily long period
of development compared with other species. Throughout this period, up to and including our teenage years, the human brain is being shaped and sculpted by experience in the cultural world. Menary and Gillett show how brain plasticity makes it possible for the cultural niche of humans to transform and enhance human cognitive capacities. They develop this argument through a case study of mathematical cognition, discussing in particular findings that in the human brain we can find a system that responds to approximate quantities (the so-called “ancient number system”) and a system that allows for reasoning about precise quantities (the “discrete number system”). This is an example of how culturally developed capacities repurpose phylogenetically older regions of the brain involved in processing number for newer cultural purposes.

Collaborative and cooperative social interactions are ubiquitous in human social life but these interactions tend to be selective. People exhibit a marked preference for cooperating with those who belong to their own social group and actively work against outsiders. In chapter 5 Edouard Machery discusses the evolution of what he terms “tribe psychology”. While tribes have largely disappeared from the modern world, many social groups are organised in ways that closely resemble tribes. Tribal thinking may also have left its mark on human social psychology in the form of nationalism, ethnocentrism and racism, as Machery explores in his chapter. While it is controversial to what extent tribal psychology leads us to think of differences between groups as essential and immutable, Machery suggests this aspect of human psychology may be behind deep rooted distinctions between “us” and “them” which seem built into human moral psychology.

Machery explains how the distinction between us and them may have originated in the tribal psychology that emerged with the cooperative and collaborative lifestyle of early humans. In the closing chapter of this section, psychologist John Barresi draws upon the same period in human history to offer a genealogy of egalitarian or agent-neutral reasons and values. He shows how the cooperative and collaborative lifestyle of early humans may have led them to the development of the concept of personhood. Barresi shows how “agent-neutral” reasons and motivations are grounded in a conceptual ability people have to think of themselves and others as persons. He argues that it was the integration of first-person and third-person perspectives that occurs in joint action that may have set the stage for early humans to develop an agent-neutral perspective.

2. Developmental and comparative perspectives

The chapters in part 2 explore cognitive development in three domains essential to human social life: 1. understanding other minds; 2. morality; and 3. shared and collective intentionality. A central theme in this section is that humans make sense of the actions of themselves and others in a variety of ways, only some of which require reasoning about mental states, a capacity sometimes referred to as “mindreading”. A second theme concerns the effects of cultural experience on cognitive development. These chapters assess to what extent people growing up in different cultures differ in their cognitive capacities, and explore the implications of cultural variation where it is found. A third theme relates to the role of social interaction in human cognitive development. Social interaction in humans is characterised by intersubjective forms of shared experience that become increasingly elaborate over the course of development. These intersubjective forms of experience result in infants developing minds that are social at their very core because they are relationally constituted, a theme that a number of chapters return to in part 5.

In the opening chapter of part 2, philosopher and comparative psychologist Kristin Andrews assesses the evidence that it is mindreading that makes human cognition unique, setting humans
apart from other primates. She suggests that human folk psychology essentially concerns thinking about how people should act rather than thinking about how they do, or will act. Andrews shows how there is substantial evidence for thinking that primates likewise operate with normative expectations about each other. She assembles a rich body of evidence that suggests that chimpanzees, for example, construct models about how individuals belonging to their social groups ought to behave. Once we think of folk psychology as providing us with models of how people should act, she argues there is every reason to think that primates likewise operate with folk psychology that is no less rich than that found in humans.

Chapter 8, by developmental psychologist Hannes Rakoczy, discusses how different forms of intentionality develop in children and primates. He starts by providing a taxonomy of types of intentionality organised around distinctions between first- and second-order intentionality and between individual, shared and collective intentional states. Both infants and other animals develop first-order intentional states, but he shows how the evidence for second-order intentional states or meta-representational capacities in non-human animals is at best patchy. Shared intentional states are found whenever two or more individuals form a joint “we” attitude. Rakoczy reviews compelling evidence that both children and primates can share perceptual states in jointly attending to an object. However, whether primates can engage in cooperative action for joint or shared goals is more controversial. Rakoczy ends his chapter by discussing the development of collective intentionality that John Searle (2010) has argued to be necessary for the construction of entities that make up the social world such as norms, conventions and institutions.

Even if people do make use of a plurality of strategies and heuristics to make sense of each other, they undoubtedly also have recourse to mindreading. The next two chapters are concerned with mindreading proper and in particular belief understanding. In chapter 9, developmental psychologists Rose M. Scott, Erin Roby and Megan A. Smith review over a decade’s worth of research that shows how infants from around 15 months and from a variety of cultures are able to track false beliefs. In a violation of expectation paradigm, for instance, infants have been shown to look much longer when an agent acts in ways that conflict with what the agent ought to do given certain beliefs. It seems that infants expect an agent to act in one way because of what the agent ought to believe. When the agent acts differently, this surprises the infant and this is taken to be evidence that the infant represents the agent as having certain beliefs that conflict with the agent’s behaviour.

In chapter 10, Jane Suilin Lavelle considers to what extent people differ across cultures in their recourse to mindreading. Does folk psychology only capture our peculiarly Western understanding of people’s behaviour? Well-known findings from cultural psychology show that people from European and American cultures typically exhibit individualist patterns of thinking and reasoning, while people from East-Asian cultures tend to think and reason more holistically (Nisbett et al. 2001). Might these differences in patterns of thinking translate into differences in how people from these different cultural backgrounds go about explaining behaviour?

Developmental psychologists Jeremy I.M. Carpendale, Michael Frayn and Philip Kucharczyk outline two views of human development in chapter 11. What they call “individualist” theories claim that mental states are private to the minds of individuals in the sense of being known directly only by the persons to whom they belong. Individualist theories claim that to know the minds of others, a person must engage in either theory-based inference or simulation routines, or some combination of the two. Carpendale and colleagues defend what they describe as an “interactionist theory” of development according to which the mind of the infant takes form in development through social interaction. Anika Fiebich, Shaun Gallagher and Daniel D. Hutto also stress the interactive character of social cognition in chapter 12. They
argue against the received view that mindreading is the default method that people employ for understanding each other, and argue instead for a pluralistic theory of social cognition in which second-person narratives, character traits, and expectations based on norms, habits and conventions all do a large part of the work of delivering mutual understanding in our day-to-day interactions.16

Chapter 13, by developmental psychologists Philippe Rochat and Erin Robbins, discusses the development of moral agency, with a particular focus on the understanding of ownership and fairness in children. How do children come to develop an ethical stance from which they evaluate not only the fairness or otherwise of some distribution of goods but also to think about how goods ought to be shared? Rochat and Robbins describe development in terms of a progressive elaboration of shared intersubjective experiences. Around 21 months, children start to express self-conscious emotions like embarrassment, shame, envy and pride that concern the children’s evaluation of themselves as members of a group. At the same time, children exhibit sensitivity to group norms and begin to develop what Rochat and Robbins label an “ethical stance”, acting according to ethical principles of fairness. Rochat and Robbins focus in particular on inequity aversion: children across widely different cultures display a clear preference for proportional equity, preferring to distribute resources to each individual according to their need.

3. Mechanisms of the moral mind

Humans are distinguished from other animals by how much of our everyday behaviour is governed by normative rules and principles that tell people what is expected of them by others belonging to their groups and communities. They do so often without being written into the law and without enforcement by social institutions. People from across cultures are universally motivated to act according to norms, seemingly treating them as ultimate ends (Sripida & Stich 2007). Violations of norms bring with them reactive attitudes such as anger, condemnation and blame and serve as the basis for punishment. Sanctions for norm violation are found across cultures (Sober & Wilson 1998).17 The chapters in this section are concerned with the psychological and neural mechanisms that might explain why social norms are found across all human groups, governing a vast array of human activities. Why are people motivated to conform to norms? Does the motivation stem from a concern with reputation, which signals to others whether or not an agent is to be trusted? Could the motivation derive from a person’s more fundamental values and ideals?18 A second theme concerns the role of evaluative learning in moral cognition. Affect is the brain’s way of keeping track of the expected value of actions based on the history of past reward and may also underlie moral judgement and behaviour.19

In chapter 14 the evolutionary anthropologists Jan M. Engelmann and Christian Zeller explore what motivates people to act in compliance with moral norms. Are people motivated by a self-interested desire to protect their reputation? A person’s reputation is of the utmost importance in societies that depend on cooperation. Reputation tells us which people are trustworthy and which are not. One way to settle this question experimentally would be to observe whether people behave more morally when their behaviour is publically observed, when the chances are high that others might get to hear about their behaviour through gossip. Engelmann and Zeller review studies that show that people are more motivated to comply with norms when their doing so can be observed by others.20 However, they argue that these findings don’t necessarily establish that people’s motivations are always strategic. The presence of other people results in us automatically and spontaneously adopting their perspective. We immediately think about how other people will evaluate and judge us. This can be an
important source of information that allows us to assess to what extent our actions and decisions match up with our own ideals and values.

Chapter 15, by cognitive neuroscientist Fiery Cushman, explores the basis for non-consequentialist moral judgement. People don’t always make moral decisions with the goal in mind of maximising utility such as the welfare of others. Cushman asks whether failure to do so is a bug in systems that would otherwise function along consequentialist lines. To answer this question, Cushman appeals to neurocomputational models of decision-making and evaluative learning that also make an appearance in a number of other chapters in this part of the handbook. He suggests that non-consequentialist moral thinking might be explained by model-free systems involved in habit learning that assign value to actions based on a history of reward. Consequentialist moral thinking, on the other hand, may be accounted for by systems involved in deliberate planning of actions. These systems search through decision trees, set up on the basis of statistical associations between actions and their outcomes. Both systems have a role to play in the explanation of moral cognition.

Philosophers Bryce Huebner and Trip Glazer explore the social function of emotions in chapter 16. They focus on emotions that were selected for the role they play in helping animals to navigate the challenges of social life. For example, emotions such as guilt or shame can motivate us to comply with social norms when we might otherwise be disposed not to (Greene 2013; Frank 1988). Huebner and Glazer show how emotions can also motivate individuals not to conform to unfair or unjust social arrangements, and to imagine alternative social arrangements. At first glance this last function of motivating social resistance would seem to be hard to explain in terms of the biological functions of emotions. They address this problem by providing a detailed account of the role of affective capacities in attuning us to the social world. Affective states are never unambiguous in the meaning: fear, for instance, can look and feel very differently depending on the situation to which it is a response. Before the agent can settle on a course of action, their affective systems have to arrive at an adequate conceptualisation or construal of the situation. How exactly a person responds to a particular situation will thus depend on the conceptualisations they rely upon. Different conceptualisations of affect may open up new, hitherto unforeseen possibilities for social engagement.

Continuing with the theme of the role of emotions in moral behaviour, chapter 17, by philosopher Michael Brownstein, asks under what conditions intuition and emotional reactions can give an agent credible reasons for action, guiding the agent’s actions in ways that fit with their rational judgement. There is substantial evidence that people navigate the social world based on implicit attitudes, many of which are learned through processes of model-free learning. By attuning us to the social world, implicit attitudes possess a defeasible moral credibility. What should we say, however, when these learning systems are tuned up by an environment suffused with prejudice? How can a person recognise when his morally credible spontaneous inclinations have become corrupted by living in an unjust social world? Brownstein recommends that we think of the cultivation of the moral implicit attitudes as a type of skill learning. Still, he notes, this leaves as an urgent and open question what social skills are required for steering a moral path through an environment dominated by prejudice and injustice.

Chapter 18 by Matteo Colombo brings this section to a close by returning to the topic of social motivation. His chapter explores to what extent the empirical literature in computational neuroscience confirms Humean theories that claim that beliefs have no motivational power taken in isolation from desires. Colombo begins by reviewing computational models of learning that support Humeanism by distinguishing mechanisms that compute value or utility from mechanisms that compute probabilities. Predictive processing theories of the brain, however, reject such a distinction, modelling all neural processing as working in the service of

Julian Kiverstein
the accurate and precise prediction of current sensory input. Colombo finishes his chapter by asking what the latter view might imply for theories of social motivation. One consequence is that neuroscientists would no longer need to invoke the concept of utility maximisation in explanations of decision-making. They could instead frame their explanations in terms of prior beliefs about the expected sensory consequences of actions.

4. Naturalistic approaches to shared and collective intentionality

Although joint action is seen in many species of animal (e.g. bees swarming, chimpanzees hunting together), humans seem to be motivated to coordinate their actions for joint goals far more flexibly and in a wider range of contexts than is seen in other animals. Capacities for joint action and joint attentional engagement allow people to coordinate with each other in ways that set the stage for agreeing upon conventions and norms. What is it for two or more people to have intentional states (e.g. plans) that are jointly directed at the world? Does it require people to have special capacities for forming group mental states such as joint commitments (Gilbert 2013) or we-intentions (Searle 1995) that are not explained by the psychological capacities of individuals? The chapters in this part of the handbook all share in common the aim of offering naturalistic or psychologically informed accounts of shared and collective intentionality. The section begins with discussions of shared intentionality in the context of joint action. These chapters develop a minimalist theory of joint action that offers an account of joint action without invoking special collective forms of intentionality. An urgent question for such a minimalist approach is how to account for the difference between thinking in the first-person singular (or I-mode) and thinking and reasoning in the first-person plural (or we-mode).

A further set of issues tackled in this section concern how people reason in the we-mode in solving coordination problems (Karpus & Gold, ch.23; Chater et al., ch.24). Coordination problems arise in the context of cooperative actions in which knowing what the other person will do is crucial to the success of the project. My decision in this kind of strategic interaction is conditional on knowing what you will do, and your decision is likewise dependent on knowing what I will do. Solving this problem is particularly crucial since norms, conventions and institutions are arguably solutions to coordination problems. If we are to understand how people came to organise society around conventions and institutions, we need to understand how they solve coordination problems.

Chapter 19 by Angelica Kaufmann opens this section with a critical evaluation of the thesis defended by Tomasello and colleagues that shared intentionality is unique to human primates (Tomasello 2014). Her chapter is focused on evaluating this thesis for the capacity to form and coordinate joint plans of action. Kaufmann focuses on the examples of group hunting as observed in Tai chimpanzees’ group hunting behaviour by cognitive ethologist Christophe Boesch. She argues that group hunting is a planned joint activity, a claim she defends by relying on Bratman’s recent account of shared intention (Bratman 2014).

Chapter 20 by Stephen Butterfill continues where Kaufmann leaves off. Butterfill defends a “minimalist” account of joint action that avoids introducing special types of plural subjects and does not attribute a special class of attitudes to agents such as we-intentions or joint-commitments. Butterfill addresses a central question that has occupied philosophers’ attention in the literature on shared intentionality: what distinguishes actions that are performed by agents in parallel from genuinely joint actions? A minimalist answer to this question looks for the simplest possible cases of joint action and only adds extra ingredients as and when they are strictly necessary. Take as an example of a minimal case two people using a system of a pulley
Julian Kiverstein

and rope to move a heavy block. Butterfill describes lifting the block as the “collective goal” of their actions. It is a collective goal in the sense that it is a goal that can only be brought about through the coordinated actions of the two agents. Butterfill shows how the concept of collective goals yields an account of joint action that is too broad. He therefore proposes to supplement it with a requirement that each of the agents that takes part in a joint action acts for an intention fulfilled by actions that have collective goals. Butterfill labels this “the flat intention” view of joint action because it avoids appeal to higher-order intentions (i.e. intentions that agents intend to act together).

Chapter 21 by philosopher John Michael constructs a minimalist account of commitment in the context of joint action. On standard views of commitment agents must form conditional obligations that are common knowledge between agents. Yet children and non-humans show an understanding of commitment without understanding much if anything about obligations that arise from promises (Gräfenhain et al. 2009). Michael develops a minimalist theory of commitment in order to address this and other problems. He argues that agents have default expectations that others will help them to bring about goals that they cannot bring about on their own. We see this in infants, for instance, that expect others to help them attain desires they are not able to fulfill by themselves. Such an expectation automatically triggers a sense of commitment to help the child to bring about the goal she cannot achieve by herself. A minimal theory of commitment can therefore make sense of how animals and infants can form commitments without understanding anything of acts of promising and the obligations such acts imply.

In chapter 22 philosopher Mattia Gallotti turns to more general questions concerning the nature of collective intentionality. Gallotti asks what it is about the attitudes of a plurality of individuals that make it the case that they are shared, or jointly directed at the world. In common with previous chapters, Gallotti resists explaining this fact in terms of metaphysical entities such as plural subjects or group minds. The psychological states needed for understanding collective intentionality belong to individuals but occur only in the context of group interactions. It is for this reason that Gallotti describes collective intentionality in terms of the “first-person plural” or the we-mode. He draws on research from the cognitive sciences to argue that the we-mode can be explained at the sub-personal level in terms of processes of mutual or reciprocal alignment (Tollefson & Dale 2012).

The next two chapters take up the question of how people succeed in solving coordination problems. Chapter 23, by philosophers Jurgis Karpus and Natalie Gold, discusses strategic decision-making in the Hi-Lo and prisoner’s dilemma games. In these games the player’s decision about what to do is conditional on beliefs about what other players in the game will do. Each player seeks the optimal strategy that will deliver the highest pay-off, but in order to select the optimal strategy players have to know what the other player will decide. Karpus and Gold show how team reasoning can help us to solve this problem. We think about what outcome would serve the interests of both of us the best, and then select the strategy that allows each of us to play our part in bringing about this outcome.

Chapter 24, by behavioural scientists Nick Chater, Jennifer Misyak, Tigran Melkonyan and Hossam Zeitoun, also discusses how people use group or we-thinking to solve coordination problems. Chater and colleagues propose an account of group-thinking in terms of what they call virtual bargaining. We think to ourselves: what would we both agree upon if we were to discuss the problem at hand? The difference between virtual bargaining and team reasoning is that it doesn’t just apply to cooperative games in which there is some common goal two or more agents aim to bring about. Chater and colleagues claim virtual bargaining can also be used to solve decision problems in which two or more individuals have competing interests, and each has an interest in double-crossing the other. Chater and colleagues go on to show
Introduction

how virtual bargaining can also explain how coordination problems are solved in communication and in joint action. In each of these domains, answering the question of what I should do is conditional on my knowing what the people I am interacting with are likely to do.

Part 4 closes with chapter 25 by Ron Mallon, which investigates the mechanisms that underlie socially constructed categories such as sexuality, gender and race. Mallon begins by asking whether these categories can be thought of as social roles that people play. Social roles confer on groups of people certain rights, duties and expectations that are common knowledge among people that understand these social categories. However, race, gender and sexuality are unlike other social role categories in an important respect. People treat the latter categories and the differences they mark as if they were natural kinds that identified essential differences between people. Mallon calls them “covert constructions” because unlike other socially constructed categories, people mistake race, gender and sexuality for natural kinds. Covert constructions are sustained and continue to exert a causal influence on people’s behaviour because of people’s mistaken belief in their naturalness.

5. Social forms of selfhood and mindedness

The chapters in this final part of the handbook explore arguments that there are aspects of the mind that the individual person has only by virtue of being a member of a social group. The chapters by Jesse J. Prinz and Shaun Nichols (ch.26) and by Mark Alfano and Joshua August Skorburg (ch.27) both take up the relationship between a person’s moral identity and the social environment. Prinz and Nichols discuss the concept of personal identity and argue that a person’s identity depends upon the links a person forges with others as a part of a moral community. Alfano and Skorburg continue this line of thought through a discussion of epistemic and moral virtue and vice. They show how virtue and vice can depend with different degrees of strength on the social environment.

Self-knowledge has often been associated with a peculiar epistemic authority or privilege that doesn’t hold in the case of knowledge of the mental states of others. People know their own minds immediately and directly and the minds of others only on the basis of observations of their behaviour. The next three chapters complicate this received view on the relation between self and other. Tadeusz W. Zawidzki (ch.28) focuses on self-interpretation and the role it plays in licensing expectations from other people. He makes an argument for understanding self-knowledge in a social context. The chapters by Pierre Jacob and Frédérique de Vignemont (ch.29) and Dan Zahavi and Alessandro Salice (ch.30) both discuss empathy and the access it gives us to another person’s emotional life. Jacob and de Vignemont focus on the sense in which empathy allows us to live through another person’s emotional experience as if it were our own. Zahavi and Salice, by contrast, stress that empathy delivers direct perceptual awareness of an experience that is not our own. The final two chapters by Glenda Satne (ch.31) and Joseph Rouse (ch.32) take up arguments that the human mind is constitutively social. Satne focuses on arguments for this claim based on the social nature of the intentional relation, while Rouse focuses his attention on arguments for the normativity of meaning.

Chapter 26, by Jesse J. Prinz and Shaun Nichols, opens the section. Prinz and Nichols present previously unpublished experimental findings about the kinds of continuity people take to be important for the identity of a person over time. They describe a series of experiments they have carried out in which they ask participants to consider whether a person was still the same person after undergoing a moral change due to a brain injury. They compared people’s responses with cases in which a person experiences loss of memory, agency and narrative. They found that people tended to rate continuity in moral values as more important than continuity
of these other cognitive capacities. Thus retaining the same moral values seems to be a central component in what people ordinarily believe makes someone the same person as they undergo changes in their lives. A person’s moral conduct in relation to others is rated by people as an important determinant of a person’s identity through time.

In Chapter 27 Mark Alfano and Joshua Skorburg argue for the dependence of moral character on the social environment. The chapter takes as its starting point much-discussed findings from social psychology that show that people’s moral behaviour can be influenced by all manner of environmental factors that ought really to have no influence on virtuous moral agents. Social expectations seem to be a particularly powerful predictor of moral behaviour. Alfano and Skorburg argue for an interpretation of these situational influences as pointing to dependence of character on the social environment. We act as the people around us expect us to. Character is not a monadic property of individual agents, but strongly depends on social relations.

The next three chapters take up questions concerning self-knowledge and the nature of knowledge of other minds. In Chapter 28 Tadeusz W. Zawidzki argues that this epistemic privilege derives from the ability of humans to shape each other’s thoughts and actions in ways that make us more predictable to each other. Zawidzki proposes that self-attributions of propositional attitudes can be thought of as commitment devices that provide an agent with an incentive to act in ways that conform to what he has said about himself. The role of the self-attribution is thus to make the person committed to being the kind of person that conforms to what others expect from them given this self-attribution. Zawidzki argues that this is a role that self-attributions can only play for the self since we can speak for ourselves with an authority that we cannot speak for others.

Chapter 29, by Pierre Jacob and Frédérique de Vignemont, is concerned with knowledge of other minds, in particular knowledge of emotions and sensations that we get through empathy. Jacob and Vignemont take empathy to be a vicarious experience in which a person undergoes an affective experience similar to an affective experience belonging to the target of their empathic episode. For example, if x is a child scared of bullying at school and y empathises with x, then it is necessary for y also to be vicariously afraid of bullies. Jacob and Vignemont take empathy to depend upon mental simulation, which they analyse as a non-propositional form of imagination. When I empathise with a child’s fear, I activate my own fear system in offline simulation mode. Jacob and Vignemont argue that the vicarious experience I come to enjoy is the result of my sharing with the child an evaluative representation with the same content as that of the child (i.e. the danger of some sort of threatening stimulus).

Chapter 30, by Dan Zahavi and Alessandro Salice, shows how the phenomenological tradition in philosophy has rich insights to offer concerning the foundations of sociality. Zahavi and Salice deny that empathy requires shared experience. They argue for the reverse claim that any shared experience requires empathy. If you and I share an experience of joy, for instance, we undergo an experience in which the joy is experienced as ours, not only as yours and mine. We could not experience the joy as our joy were it not the case that each of us was already aware of the other’s joy. At the same time, empathy requires self-other differentiation, and this raises a puzzle that Zahavi and Salice go on to trace through a number of thinkers. How, given such differentiation between subjects, are shared experiences nevertheless possible? Shared experience requires an interlocking or interdependence of the experiences of different subjects. Zahavi and Salice discuss a number of proposals for how to characterise this interdependence.

Chapter 31, by Glenda Satne, critically evaluates two theories that take intentionality to, in some way, depend on social practices. Communitarian accounts of intentionality analyse meaning by reference to consensus among members of a community about how a term or
concept ought to be used. Individuals are trained through mechanisms of social conformism to use a concept or term as others do. Interpretivist theories of meaning take intentionality to originate in a conceptual framework people employ to make sense of each other as rational agents. Satne shows how both these approaches face a number of difficulties, not least of all in accounting for the development of intentional capacities. In order for these accounts of meaning to get off the ground, it seems they have to assume that agents are already in possession of intentional capacities. She ends her chapter by arguing that this debt can be repaid by enactivist theories that argue for a theory of basic cognition as content-free (Hutto & Satne 2015).

In the final chapter (ch.32), Joseph Rouse takes up the question of whether mind and meaning are normative phenomena. Rouse defines normativity as phenomena the response to which can be understood as correct or incorrect, appropriate or inappropriate, right or wrong, just or unjust and so on. He then goes on to explore arguments that have been made that discursive or conceptual capacities might be thought of as normative in this sense. A normative theory of these capacities claims that to think, speak or act is to be answerable to what one ought to think, say, or do. Normative theories face a problem, however, when it comes to situating normative phenomena in nature as understood scientifically. The response to this problem developed by Rouse is to ground normativity in human social life. In agreement with Satne, Rouse argues that both communitarian and interpretivist versions of this strategy face insuperable difficulties. Rouse's own preferred view locates the normativity of social practices in the temporality of human agency, whereby members of a practice are oriented to a common future on the basis of their present and past performances.

Conclusion

The essays in this handbook are unified by their ambition to provide a naturalistic account of the sociality of the human mind. These naturalistic projects take very different forms, a flavour of which I have attempted to impart in providing an overview of the volume as a whole. Taking seriously the sociality of the human mind has two important consequences. First, it directs our attention to phenomena that normally fall within the remit of the social sciences and might not otherwise receive the attention of philosophers and cognitive scientists concerned with the nature of the human mind. Second, it corrects for an individualist orientation both in philosophy and in the cognitive sciences. Most philosophers and cognitive scientists concerned tend to think of the individual person as the locus of mental states. The essays assembled in this volume aim to show that membership of social groups, and the coordination that takes place in social interactions, are just as important for understanding the minds of humans as the processes taking place inside of individuals’ heads. Indeed, understanding the fundamental sociality of mind may even help people to imagine new social arrangements that correct for the injustices and imbalances of power inherent in the societies we live in today.30

Notes

1 Machery (ch.5) applies this approach to ethnic cognition, and the chapters by Scott et al. (ch.9) and Lavelle (ch.10) to theory of mind. See Barrett (ch.1) and Carpendale et al. (ch.11) for critical discussions of evolutionary psychology.

2 Byrne and Whiten (1989) further developed this hypothesis showing how the demands of primate social life may have led to the development of complex capacities for reasoning about the psychological states of conspecifics. Dunbar (1995) argues that selection pressures arising from group living led to increases in brain size. In particular it was the challenge of keeping track of and manipulating information about large numbers of pairs of adult mates that proved cognitively demanding.
Intentionality is used here to refer to the feature of mental states in virtue of which they are directed towards objects and states of affairs in the world (see Rakoczy (ch.8) and Satne (ch.31)). Collective and shared intentionality occurs when two or more individuals have mental states that are jointly directed at objects or states of affairs. Collective and shared forms of intentionality are a central topic in this volume and are discussed throughout part 4.

This claim that humans are unique in their capacity for shared intentionality is discussed by Barresi (ch.6) and Rakoczy (ch.8). See the chapters by Andrews (ch.7) and Kaufmann (ch.19) for critique.

Zawidzki (ch.28) argues that cooperative communication may well have also originated through selection pressures for the better and more complex forms of signalling of cooperative potential.

See the chapters by Moore (ch.3), Boutel and Lewens (ch.2), Menary and Gillett (ch.4), Barresi (ch.6), and Zawidzki (ch.28).

See Menary and Gillett (ch.4) for further discussion.

This claim, that humans are unique in their capacity for shared intentionality, is controversial. See Andrews (ch.7) and Kaufmann (ch.19) for evidence against this claim.

See Mallon (ch.25) for further discussion of essentialist thinking and its role in thinking about gender, sexuality and race.

See the chapters by Andrews (ch.7), Carpendale et al. (ch.11) and Fiebich et al. (ch.12) for further discussion.

See the chapters by Scott et al. (ch.9), Lavelle (ch.10), Fiebich et al. (ch.12) and Rochat and Robbins (ch.13).

See the chapters by Carpendale et al. (ch.11), Fiebich et al. (ch.12) and Rochat and Robbins (ch.13).

See in particular the chapter by Zahavi and Salice (ch.30).

Shared intentionality is among the topics explored in part 4 of this handbook. See in particular the essays by Kaufmann (ch.19), Butterfill (ch.20) and Gallotti (ch.22).

This is a question that is taken up by philosopher Angelica Kaufmann in detail in Chapter 19.

There is an interesting agreement with the arguments of Andrews (ch.7) here.

The chapter by Rochat and Robbins (ch.13), for instance, shows how children across cultures are motivated by fairness norms.

See the chapters by Englemann and Zeller (ch.14), Huebner and Glazer (ch.16) and Colombo (ch.18) for discussions of norm compliance.

The chapters by Cushman (ch.15), Huebner and Glazer (ch.16) and Brownstein (ch.17) show how processes of non-social reward learning play a central role in moral cognition. Colombo (ch.18) also discusses these models, contrasting them with the account of motivation that is supported by predictive processing models of decision-making.

They focus in particular on a study in which people have to decide whether to pay into an honesty box for drinks they are consuming in an office canteen (Bateson et al. 2006). When a picture of eyes was placed next to the instructions for the honesty box, Bateson and her colleagues found that nearly three times more people paid as compared with what happened when a picture of flowers was placed by the instructions. The eyes seem to function as a cue that reminds the agent of the reputational costs of freeriding.

For an excellent overview of the predictive processing theoretical framework in computational neuroscience, see Hohwy (2013).

See the chapters by Kaufmann (ch.19), Butterfill (ch.20) and Michael (ch.21).

See the chapters by Gallotti (ch.22), Karpus and Gold (ch.23) and Chater et al. (ch.24).

For instance, in the Hi-Lo game players must coordinate on the same response in order to obtain pay-offs. If they pick options that do not match, neither of them receives anything. In order to receive the highest pay-off, they must both pick Hi, but how can I know for sure that if I pick “Hi” you will do the same? (Bacharach 2006)

The chapters by Karpus and Gold (ch.23), Chater et al. (ch.24) and Mallon (ch.25) explore to what extent theories of reasoning in the we-mode and collective intentionality can help us to answer this question and these more general questions about the construction of social reality.

Karpus and Gold state in their chapter that double-crossing decreases the likelihood that people will engage in team-reasoning. Furthermore, Chater and colleagues show that if people with competing interests were to team-reason they would agree on a sub-optimal strategy, whereas if they were to engage in virtual bargaining this would result in them selecting a strategy with the best pay-off for both of them.

Zawidzki is following Brandom 1994. For critical discussion of Brandom, see Satne (ch.30) and Rouse (ch.31).
Introduction

28 They therefore disagree strongly with the account of empathy as vicarious experience defended by Jacob and Vignemont in the previous chapter. Jacob and Vignemont, for their part, offer a number of arguments against the phenomenological theories in the opening section of their chapter (ch.29).

29 I use “intentionality” as a technical term to refer to the meaningful directedness of the mind at the world.

30 See the chapters by Huebner and Glazer (ch.16), Brownstein (ch.17) and Mallon (ch.25) for more discussion of the political implications of thinking through the sociality of the human mind.

References


